基于改进YOLOv8n-pose和三维点云分析的蒙古马体尺自动测量方法

李明煌[1,2]，苏力德[1,2]，张 永[1,2*]，宗哲英[1,2]，张 顺[3]

（1. 内蒙古农业大学 机电工程学院，内蒙古呼和浩特 010018，中国；2. 牧草饲料生产全程智能化装备内蒙古自治区工程研究中心，内蒙古呼和浩特 010018，中国；3. 内蒙古自治区巴彦淖尔市现代农牧事业发展中心，内蒙古巴彦淖尔 015001，中国）

摘 要：［**目的/意义**］准确高效地获取马匹体尺信息是马产业现代化进程中的关键环节。传统的人工测量方法耗时长、工作量大，且会对马匹造成一定应激反应。因此，实现准确且高效的体尺参数自动测量对于制定蒙古马早期育种计划至关重要。［**方法**］选择Azure Kinect深度相机获取蒙古马双侧RGB-D数据，以YOLOv8n-pose为基础，通过在C2f模块中引入可变形卷积（Deformable Convolution v2, DCNv2），同时添加洗牌注意力机制（Shuffle Attention, SA）模块和优化损失函数（SCYLLA-IoU Loss, SIoU）的方法，利用余弦退火法动态调整学习率，提出一种名为DSS-YOLO（DCNv2-SA-SIoU-YOLO）的模型用于蒙古马体尺关键点的检测。其次，将RGB图中的二维关键点坐标与深度图中对应深度值相结合，得到关键点三维坐标，并实现蒙古马点云信息的转换。利用直通滤波、随机抽样一致性（Random Sample Consensus, RANSAC）、统计离群值滤波、主成分分析法（Principal Component Analysis, PCA）完成点云处理与分析。最终根据关键点坐标自动计算体高、体斜长、臀高、胸围和臀围5项体尺参数。［**结果和讨论**］DSS-YOLO的平均关键点检测精度为92.5%；d_{DSS}为7.2个像素；参数量和运算量分别仅为3.48 M和9.1 G。体尺参数自动测量结果与人工测量值相比，各项体尺参数的整体平均绝对误差为3.77 cm；平均相对误差为2.29%。［**结论**］研究结果可为蒙古马运动性能相关遗传参数的确定提供技术支撑。

关键词：蒙古马；体尺测量；卷积神经网络；注意力机制；三维点云处理；YOLOv8n-pose

中图分类号：TP391；S823 **文献标志码：**A **文章编号：**SA202312027

引用格式：李明煌, 苏力德, 张永, 宗哲英, 张顺. 基于改进YOLOv8n-pose和三维点云分析的蒙古马体尺自动测量方法[J]. 智慧农业(中英文), 2024, 6(4): 91-102. DOI：10.12133/j.smartag.SA202312027

LI Minghuang, SU Lide, ZHANG Yong, ZONG Zheying, ZHANG Shun. Automatic Measurement of Mongolian Horse Body Based on Improved YOLOv8n-pose and 3D Point Cloud Analysis[J]. Smart Agriculture, 2024, 6(4): 91-102. DOI：10.12133/j.smartag.SA202312027 (in Chinese with English abstract)

0 引 言

内蒙古自治区被誉为蒙古马故乡，有着悠久的养马、育马及赛马等文化传统，拥有发展现代马业深厚的文化底蕴和优越的地理环境与气候条件。蒙古马作为内蒙古自治区马匹的优良品种已有上千年的历史，一直是中国的一种重要资源，被农业农村部列为国家级畜禽遗传资源保护名录之一[1]。近年来，内蒙古传统养马业正在逐渐朝着集文化、经济、休闲于一体的新型第三产业发展，对于地方马品种资源有效保护、开展优良选育工作以及充分挖掘优良种质特征已经成为了发展过程中的关键环节[2]。动物育种的中心任务是对数量性状的改良，其中体尺参数在蒙古马育种工作中具有重要地位，能够直接反映其生长发育情况和品种改良效果[3,4]。体尺参数的不同可能导致运动学特征参数及动力学

收稿日期：2023-12-28

基金项目：国家自然科学基金（32360856）；内蒙古自治区自然科学基金（2022QN03019）；内蒙古自治区高等学校创新团队（NMGIRT2312）；内蒙古农业大学高层次人才引进科研启动项目（NDYB2020-21）

作者简介：李明煌，研究方向为机电控制系统及智能技术。E-mail：a1286862045@163.com

*通信作者：张 永，博士，教授，研究方向为数字化装备及自动控制技术。E-mail：yongz@imau.edu.cn

参数的变化，进而影响马匹运动性能[5]。然而传统的体尺参数测量方法主要通过手动测量完成。该方法不仅效率低、工作量大、自动化程度不高而且易造成马匹的应激反应[6]。而快速、精准、高效的测量方法是突破这一困境的关键手段。

随着现代测试技术不断发展，运用先进技术获取马匹体尺参数并评估其形态特征，以制定育种计划成为当下的主流研究方向[7-9]。Kristjansson等[10]对影响冰岛马的骑行能力的不同形态特征的遗传关系进行评估，结果表明，可通过育种方式改进体型性状以达到提高冰岛马的骑行能力。Pallottino等[11]使用立体视觉系统，通过手动方式测量利比赞马的体尺参数，结果显示测量误差小于3%。Gmel等[12]使用tpsDig2图像分析工具由操作人员手动测量提取弗赖伯格马的各项体尺参数，并采用限制性最大似然模型估计体尺参数的遗传率。Pérez-Ruiz等[13]采用LiDAR（Light Laser Detection and Ranging）传感器记录安达卢西亚马侧面信息并进行三维重建操作，通过手动缩放的方式测量体尺参数，得到人工测量值与三维重建测量值的平均相关系数为0.55。Freitag等[14]研究在美国夸特马左侧放置标记物，并结合ImageJ图像分析软件，通过人工标记的方式测量相应的体尺参数，测量结果平均相对误差小于1.5%。上述研究表明，通过图像处理方法获取体尺参数并用于马匹育种计划的制定是可行的。然而，上述方法主要依赖于人工手动选择测量点，过程复杂且自动化程度低。快速、精准、高效的测量方法已成为推进马匹育种进程的关键手段。

近年来，得益于人体姿态估计领域的进展，国内外学者将基于卷积神经网络的关键点检测方法应用于家畜关键点的实时自动检测研究中，并取得了出色的成果。Li和Teng[15]利用Hourglass模型定位奶牛和山羊躯干分割图像中的测量点位置。Wang等[16]通过构建HRNet（High-Resolution Network）模型实现检测生猪站立时的关节点。Zheng等[17]以活种鸡为试验对象，提出种鸡小腿长周长测量算法（Shank Length and Circumference Measurement, SL-CM）。该算法首先利用YOLOv5s（You Only Look Once）模型检测鸡爪和小腿根部区域，其次定位区域中的关键点，进而计算种鸡大腿体尺参数。赵宇亮等[18]利用DeepLabCut工具箱实现生猪背部点云中测量点的检测定位，从而完成对应体尺测量。

上述研究中，家畜体尺测量多是依赖单侧图像数据，胸围等重要三维体尺参数无法获取。此外，关键点检测模型复杂程度较高、泛化性不足，在养殖环境中较难实践应用。针对以上问题，本研究对YOLOv8n-pose模型进行改进，提出一种新的检测模型DSS-YOLO（DCNv2-SA-SIoU-YOLO）关键点检测模型，并结合点云数据处理方法，通过检测蒙古马三维点云中测量点位置，实现体高、体斜长、臀高、胸围和臀围5项体尺参数的自动测量，为高效、精准的蒙古马体尺测量方案提供技术支持和参考依据。

1 材料与方法

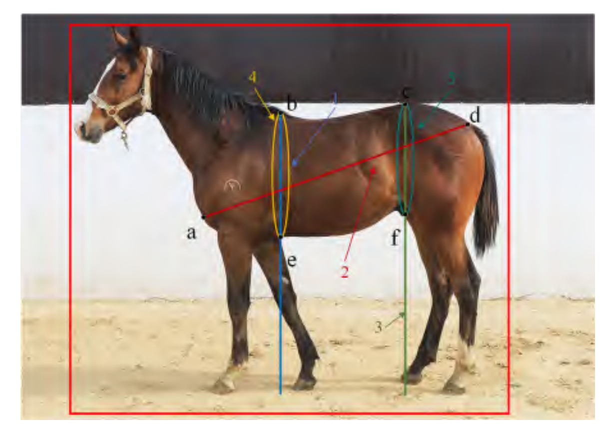
1.1 数据集构建

本研究数据分别于2022年10月和2023年4月、9月在内蒙古农业大学马匹养殖技术中心处采集。图像数据采集设备为Azure Kinect深度相机（RGB图分辨率为1 920×1 080像素；深度图分辨率为640×576像素；帧率为30 FPS），采集对象为处于适宜育种阶段的成年蒙古马，年龄（7±1.5）岁，体重（485±50） kg。拍摄处于不同场景、时间段下的行走或站立的单匹蒙古马完整侧视RGB图像视频，共采集50匹蒙古马，每段视频时长为15～25 s，视频保存为*.mp4格式，用于构建训练集和验证集，以完成DSS-YOLO关键点检测模型的训练及性能评估。另外采集30匹蒙古马RGB-D图像用于构建测试集，以完成体尺参数测量任务。

1.1.1 训练集和验证集构建

将采集到的50段RGB图像视频设置10帧/s抽取处理，并经人工筛选剔除重复、严重遮挡等质量不佳图像，共获取2 500张高清的蒙古马RGB图像，其中处于运动行走和休息站立状态的马匹图像数量比例为1∶1。为了保证模型的泛化性，并考虑实际图像捕捉过程中存在角度、光线和图像模糊的问题，利用几何变换、亮度调节以及噪声添加对原始数据集进行增强操作，最终生成包含4 000幅图像的蒙古马图像数据集，用于DSS-YOLO模型的训练及性能评估。按照8∶2的比例将标注完成的数据集分为训练集及验证集，其中训练集和验证集分别为3 200张和800张。利用Labelme图像标注软

件[19]，以蒙古马的各个测量点为参考对数据集进行标注，如图1所示。



注：1. 体高；2. 体斜长；3. 臀高；4. 胸围；5. 臀围；a. 肩部端点；b. 鬐甲点；c. 臀部最高点；d. 臀部端点；e. 肘部端点；f. 跗关节点。

图1　蒙古马体尺参数测量图示及标注

Fig. 1 Diagram and illustration of Mongolian horse body sizes measurement

1.1.2　测试集构建

在试验场景中，采集30匹蒙古马用于体尺参数测量任务。由于训练员每日会牵引一定数量的蒙古马往返于马厩和训练场地，为了在不干预养殖场日常工作的前提下进行数据采集，本研究在马厩内通道两侧各安装一台Azure Kinect深度相机，在每匹马往返时经过相机前进行采集，共获得60组RGB-D图像数据。为了减少拍摄过程中相机的抖动现象，利用三角支架对其进行了固定，蒙古马深度信息采集平台布局如图2所示。采集完成后由专业人员利用卷尺、测杖等工具测量30匹马的体高、体斜长、臀高、胸围和臀围5项参数。每项体尺参数共测3次，取均值作为人工测量值。

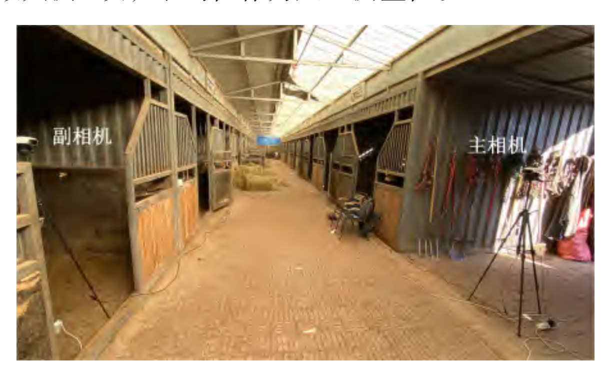


图2　蒙古马深度信息采集平台布置

Fig. 2 Platform layout of Mongolian horse deep information collection

1.2 总体技术路线图

总体技术路线如图3所示，主要由DSS-YOLO关键点检测模型和三维点云处理模块组成。首先，搭建双侧图像采集试验平台，采集蒙古马RGB图像和深度图像。其次，运用DSS-YOLO检测体尺测量的关键点信息并映射至同步采集的深度图中，得到关键点深度信息，并对深度图像进行点云信息转换，以获取蒙古马三维点云数据。最终，通过标定信息完成蒙古马双侧点云的配准和后续处理操作，并根据关键点计算各项所需体尺参数。

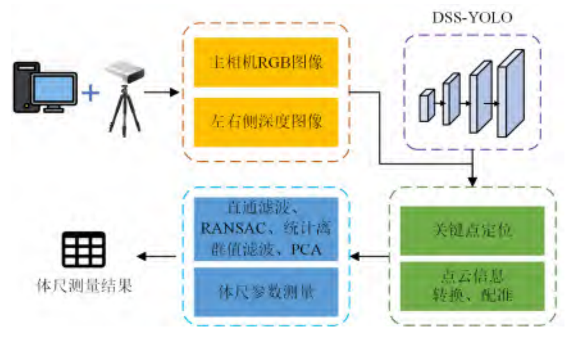


图3　蒙古马体尺参数自动测量研究技术路线图

Fig. 3 Overall technology roadmap for automated measurement of body size parameters of Mongolian horses

1.3 关键点检测模型的改进

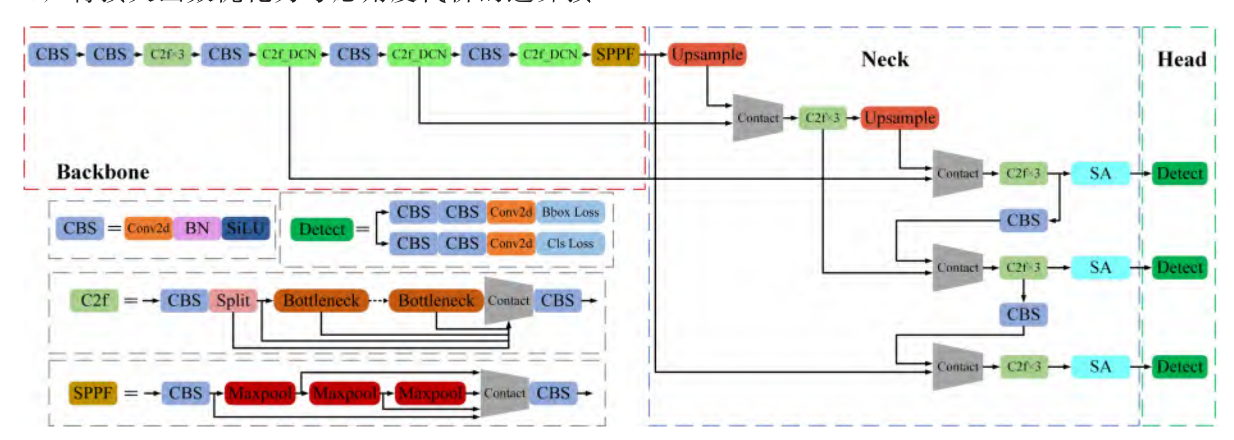
YOLOv8是YOLO系列中最新版本，相较于前代版本，该模型拥有最高的检测精度及速度[20]。此外，YOLOv8推出YOLOv8-pose单阶段关键点检测模型，网络结构主要包括输入端、Backbone主干网络、Neck颈部网络以及检测头[21]。其中，YOLOv8n-pose因其较小的计算资源和模型内存需求而具备轻量级特性，故本研究选用YOLOv8n-pose作为基础模型。由于初始YOLOv8n-pose是针对人体关节点检测定位从而实现行为识别任务的预测模型，考虑自制蒙古马关键点数据集的特性，在除关键点数目的变化之外，拟从以下3个方面对YOLOv8n-pose进行改进，以应对识别对象特征的差异，具体改进如下。

1）在原始YOLOv8n-pose模型的Backbone网络中，将C2f模块与可变形卷积（Deformable Convolution v2, DCNv2）融合，设计C2f_DCN模块结构，以帮助模型能够更好地捕捉在不同姿势下的马匹形状信息。

2）在Neck网络中添加SA洗牌注意力机制

(Shuffle Attention, SA)注意力机制模块，通过在通道信息和空间信息两个层面提取特征，能有效提高模型特征表达能力，抑制噪声特征，从而提升检测的准确性。

3）将损失函数优化为考虑角度代价的边界损失函数（SCYLLA-IoU, SIoU），加快模型训练时的收敛速度的同时，增强模型对主要图像区域的关注度，有助于降低关键点检测的错误率。改进后的网络结构如图4所示。



注：Conv2d为卷积操作；BN为批次归一化；SiLU为Sigmoid加权线性单元激活函数；Bottleneck为瓶颈结构；Contact为特征拼接；Upsample为上采样操作；Maxpool为最大池化操作；SPPF为快速空间金字塔池化。

图4 DSS-YOLO网络结构

Fig. 4 The network structure of DSS-YOLO

1.3.1 可变形卷积

传统卷积核的感受野有限，无法捕获足够的局部信息，从而影响最终检测结果。DCNv2是Zhu等[22]提出的可变形卷积网络，通过为卷积核添加不同的偏移量，从而允许卷积核在不同的采样位置添加偏移变量，可根据不同数据学习成为不同的卷积核结构，以更好捕捉目标的不规则形状和局部变化。其计算过程如公式（1）所示。

$$y(p) = \sum_{k=1}^{K} \omega_k \cdot x(p + p_k + \Delta p_k) \cdot \Delta m_k \tag{1}$$

式中：ω_k为位置k上的卷积核权重；x为输入特征映射；p_k为预先指定的偏移量；Δp_k为可学习的偏移量；Δm_k为权重系数，取值范围为［0，1］。鉴于3×3卷积核在捕捉局部特征方面的关键作用，以及Backbone网络在特征提取方面的重要性，因此本研究通过将Backbone中后3个C2f模块中的3×3卷积核替换为DCNv2，从而设计构建C2f_DCN模块[23]。改进后的网络结构如图5所示。这一改进旨在通过可变形卷积的高级特征提取能力，以帮助模型适应实际养殖环境中的不同姿势下马匹关键点的检测需求。

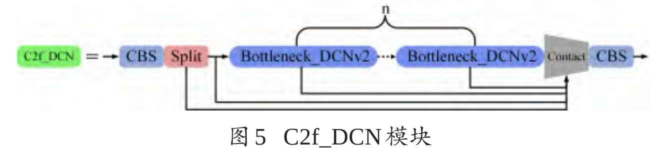


图5 C2f_DCN模块

Fig. 5 C2f_DCN module

1.3.2 注意力机制模块

注意力模块能够有效地提高卷积神经网络模型捕捉输入数据中特定区域或元素的能力，从而实现更有效、准确地预测或表征。在本研究中，通过添加SA注意力模块用于提高模型对特征信息的专注度，同时过滤冗杂信息[24]，其网络结构如图6所示。此外，SA具有高度的灵活性可以被添加到任何卷积层的末端。

本研究设计在Neck网络中的C2f卷积模块之后添加SA注意力模块。该设计可以提高模型在不同阶段对于特征信息的关注程度，进而提升关键点检测准确度。

1.3.3 损失函数

YOLOv8n-pose采用完全交并比损失（Complete IoU Loss, CIoU）和分布式聚焦损失函数（Distribution Focal Loss, DFL）[25]作为边界框损失函数来衡量预测框与真实框之间的重叠程度。但CIoU并没有明确考虑预测框的宽度和高度不能同时按一定比例收敛，从而导致预测框与真实框之间存在着方向上的差异，导致模型训练时的收敛效果较差，进一步导致训练结果不理想。针对该问题，

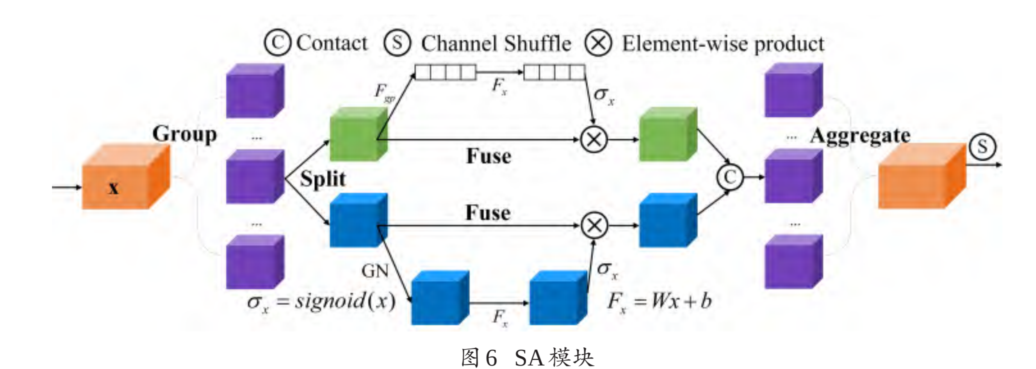


图6　SA模块

Fig. 6 SA module

本研究采用SIoU作为边界框损失函数[26]。SIoU考虑预测框与真实框之间的矢量角度、距离、形状及IoU成本，由于本研究着重于中大型目标检测，而且自定义的关键点均处在边界框内，提高边界框准确度有助于降低关键点检测错误率。定义如公式（2）所示。

$$L_{SIoU}=1-IoU+(\Omega+\Delta)/2,\quad IoU=\frac{b\cap b_{gt}}{b\cup b_{gt}} \tag{2}$$

式中：Δ为距离成本；Ω为形状成本；b为预测框；b_{gt}为真实框。角度成本示意图如图7所示，σ为真实框中心点与预测框之间的欧式距离。在训练过程中，如果向量夹角$\alpha\leq\pi/4$，损失函数则将优先沿x轴方向预测，使α最小化，相反则沿着y轴进行预测，使得β最小化。因此，SIoU函数规定了预测框向真实框逼近的方向，避免了训练过程中出现的预测框“徘徊”现象，可进一步加快模型训练中的收敛速度。

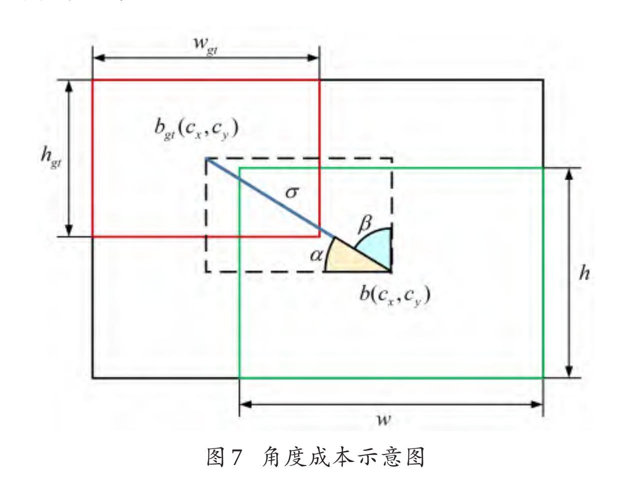


图7　角度成本示意图

Fig. 7 Schematic diagram of angle cost

1.4 三维点云转换

由于DSS-YOLO所检测得的关键点为二维坐标，需将其转换为真实环境中的三维坐标，以实现蒙古马三维点云中各项体尺参数自动测量。通过Azure Kinect SDK[27]中的校准函数，将RGB图与对应深度图进行对齐校正，并将关键点信息映射到相同时间点拍摄得到的深度图中，得到关键点深度信息[28]。随后，实现深度图中的像素坐标转换为世界坐标系下的三维坐标得到点云数据。计算过程基于公式（3）实现，具体转换步骤如图8所示。

$$Z_c\begin{bmatrix}u\\v\\1\end{bmatrix}=\begin{bmatrix}f_x&0&c_x\\0&f_y&c_y\\0&0&1\end{bmatrix}\begin{bmatrix}X_c\\Y_c\\Z_c\end{bmatrix} \tag{3}$$

式中：(u, v)为点像素坐标；(x, y)为点图像坐标；(X_c, Y_c, Z_c)为该点在相机坐标系下的三维坐标；(c_x, c_y)为图像中心坐标；(f_x, f_y)为相机焦距。

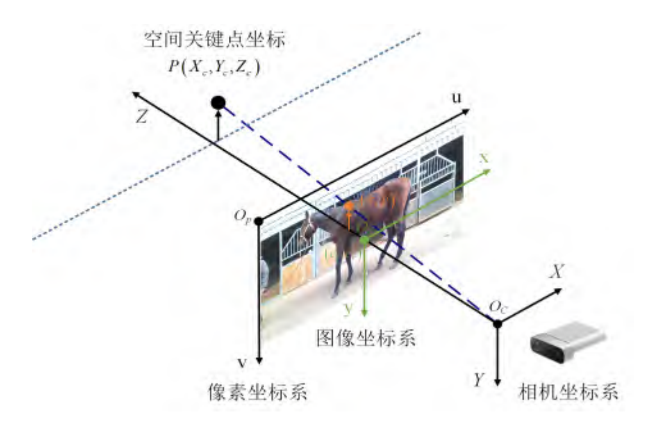


图8　三维点云坐标系转换关系示意图

Fig. 8 Three-dimensional point cloud coordinate system transformation relationship diagram

1.5 蒙古马点云处理与分析

利用立体标定后输出的两侧Azure Kinect深度相机的相对位置参数完成双侧点云配准重建。将主相机视角作为世界坐标系，通过旋转和平移副相机

生成的点云信息，得到全局点云数据，该过程基于公式（4）执行。本研究中设定主相机视角作为世界坐标系，全局点云配准效果如图9所示[29]。

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \boldsymbol{R}\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} = \begin{bmatrix} \boldsymbol{R}_1 & \boldsymbol{R}_2 & \boldsymbol{R}_3 \\ \boldsymbol{R}_4 & \boldsymbol{R}_5 & \boldsymbol{R}_6 \\ \boldsymbol{R}_7 & \boldsymbol{R}_8 & \boldsymbol{R}_9 \end{bmatrix}\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \tag{4}$$

式中：(x_1, y_1, z_1)为主相机坐标系中的一点；(x_2, y_2, z_2)为该点在副坐标系中坐标；$\boldsymbol{R}$为空间变换矩阵；$[t_x, t_y, t_z]^{\mathrm{T}}$为平移向量。



图9 实际养殖环境下的蒙古马点云全局配准效果

Fig. 9 Global alignment effect of Mongolian horse point cloud in real breeding environment

鉴于Azure Kinect深度相机的高分辨率和广角度测量距离，在获得的点云数据中存在较多噪声和无效点云数据，影响体尺参数测量精度。为确保测量准确性，本研究采取下述点云处理操作。

1）对点云配准结果使用直通滤波，在拍摄通道的3个真实坐标维度（*X*轴、*Y*轴、*Z*轴）中设置阈值，以保留目标蒙古马的点云区域。

2）由于蒙古马图像是在其运动过程中拍摄的，因此利用随机采样一致性（Random Sample Consensus, RANSAC）算法检测地平面点云位置信息并确定其对应的方程，用于体高类参数测量。拟合后的地平面点云如公式（5）所示。

$$Ax + By + Cz + D = 0 \tag{5}$$

式中：A、B、C、D为平面方程系数；x、y、z为点云坐标。

3）经直通滤波和RANSAC处理后的点云仍存在着较多噪声数据。为减少噪声点云对胸围和臀围测量精度的影响，利用统计离群值滤波去除存在于背景中及游离在体表周围的离群点云。

4）采用主成分分析法（Principal Component Analysis, PCA）对蒙古马点云进行体尺参数测量坐标系校正。将点云质心位置设为坐标原点，同时将地平面点云法向量的正方向设为*Y*轴，体长的正方向设为*X*轴，并根据右手定则，将体宽方向设为*Z*轴，以便于后续体斜长测量操作。

1.6 体尺参数测量

1.6.1 体斜长参数计算

为避免空间中点与点之间的倾斜程度影响体斜长参数的测量结果，本研究将点a和点d投影至经坐标系校正后的XOY平面上，进行体斜长参数L的计算，如公式（6）所示。

$$L = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \tag{6}$$

式中：(x_1, y_1, z_1)为点a坐标；(x_2, y_2, z_2)为点d坐标。

1.6.2 体高和臀高参数计算

根据RANSAC检测得到的地平面点云信息，计算点b和点c到地面的直线距离即为对应体高、臀高测量值，如公式（7）所示。

$$H = \frac{\left| Ax' + By' + Cz' + D \right|}{\sqrt{A^2 + B^2 + C^2}} \tag{7}$$

式中：H为体高和臀高的测量值，cm；$Ax + By + Cz + D$为地面点云方程；(x', y', z')为点b和点c的坐标。

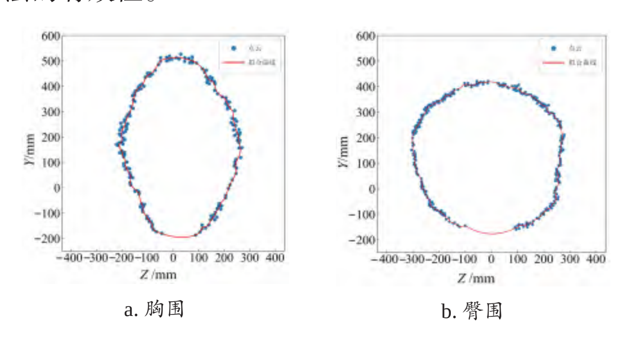
1.6.3 胸围和臀围参数计算

分别利用点b、e和点c、f的相对位置关系对点云模型做切片操作，得到胸围和臀围的轮廓点云信息。由于蒙古马胸围和臀围呈现不规则的几何特征，利用常规的圆形曲线拟合方法计算周长可能导致较大的测量误差。因此，本研究将点云切片中的各点信息转换至极坐标系中，并利用移动最小二乘法对极坐标系中连续的点云数据进行曲线拟合。随后，将拟合完成的曲线再次转换至笛卡尔坐标系下，从而得到胸围和臀围的周长信息。曲线长度S计算如公式（8）所示。

$$S = \int_{-\pi}^{\pi} \sqrt{1 + (f^{\hat{A}}(x))^2}\, dx \tag{8}$$

式中：$f(x)$为拟合后得到的曲线函数；$f^{\hat{A}}(x)$为该函数的导数。曲线拟合的结果如图10所示。由于蒙古马匹背部呈现上窄下宽的三角形态，通过对臀部点云进行剪切操作，以模拟拍摄过程可能出现的由于视角受限导致马匹背部或腹部深处点云捕获不全，从而臀部上、下侧点云会出现部分缺失的

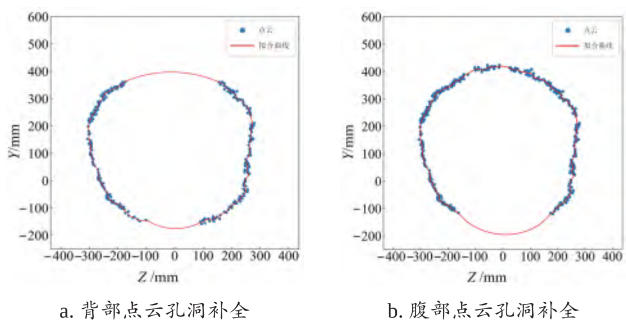
情况。剪切后的曲线拟合结果如图11所示，与理想点云情况相比，曲线拟合长度平均相对误差为2.34%，误差较小，证明了移动最小二乘拟合曲线法的有效性。



a. 胸围　　b. 臀围

图10　蒙古马胸、臀围点云切片及曲线拟合示意图

Fig. 10　Schematic diagram of point cloud slicing and curve fitting of Mongolian horse chest and croup circumference



a. 背部点云孔洞补全　　b. 腹部点云孔洞补全

图11　蒙古马背部和腹部点云出现孔洞情况时臀部曲线拟合示意图

Fig. 11　Schematic diagram of croup curve fitting with holes in the dorsal and ventral parts of the Mongolian horse point cloud

2　结果与分析

2.1　试验平台与训练参数设置

初始训练参数设置如下：初始学习率设置为0.01，学习率动量设置为0.937，输入图像大小设置为640×640，批量大小设置为16，权重衰减设置为0.000 5，优化器选择AdamW，训练轮次为150个epoch。试验平台配置参数如表1所示。模型训练过程中采用余弦退火算法动态调节学习率，帮助模型更有效地收敛，如公式（9）所示。

$$\eta_t = \eta_{min} + \frac{1}{2}(\eta_{max} - \eta_{min})[1 + \cos(\frac{T_{cur}}{T_i}\pi)] \quad (9)$$

式中：η_t为当前学习率；η_{max}为最大学习率；η_{min}为最小学习率，设置为0；T_{cur}为当前运行的epoch数；T_i为总epoch数。

表1　训练试验平台配置参数

Table 1　Configuration parameters of the training platform

配置	版本
操作系统	Ubuntu 20.04
深度学习框架	PyTorch 1.11.0
中央处理器	Intel Core i7-9700 k
图像处理器	Nvidia GeForce RTX 2080 ti
CUDA	11.3
CUDNN	8.2.1

2.2　评价指标

2.2.1　关键点检测精度评价指标

通过引入关键点相似度（Object Keypoint Similarity, OKS）计算得出关键点检测平均准确率（Average Precision, AP），其作用类似于目标检测中的IoU，当OKS大于阈值T时，结果被判别为真阳性；OKS小于等于T时，结果被判别为假阳性；OKS等于0时，结果被判别为假阴性。本研究计算在不同T（从0.5～0.95，步长为0.05）下的AP均值作为评价指标，并记为mAP，计算如公式（10）所示。

$$\mathrm{mAP} = \frac{\sum_p \delta(OKS_p > T)}{\sum_p 1} \quad (10)$$

同时，统计测试集中预测关键点与真实关键点之间平均欧式距离，以评估模型在测试集中的关键点定位精准度，如公式（11）所示。

$$d_{\mathrm{DSS}} = \frac{1}{n}\sum_{i=1}^{n}\sqrt{(x_{pi} - x_{gi})^2 + (y_{pi} - y_{gi})^2} \quad (11)$$

式中：d_{DSS}为平均误差；n为图像中关键点个数；(x_{pi}, y_{pi})和(x_{gi}, y_{gi})分别为预测点的像素坐标和人工标注点的像素坐标。

2.2.2　体尺参数测量评价指标

采用平均绝对误差（Mean Absolute Error, MAE）和平均相对误差（Mean Relative Error, MRE）作为体尺参数测量精度的评价指标，其计算如公式（12）和公式（13）所示。

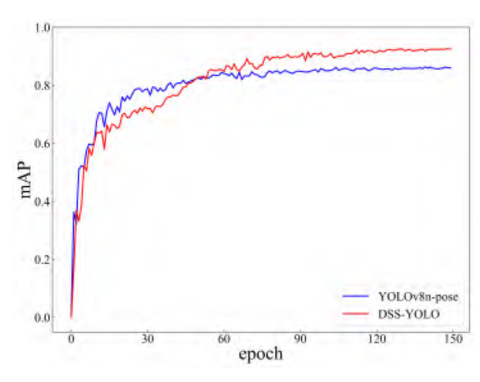
$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{x}_i - x_i\right| \quad (12)$$

$$\mathrm{MRE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{x}_i - x_i}{x_i}\right| \quad (13)$$

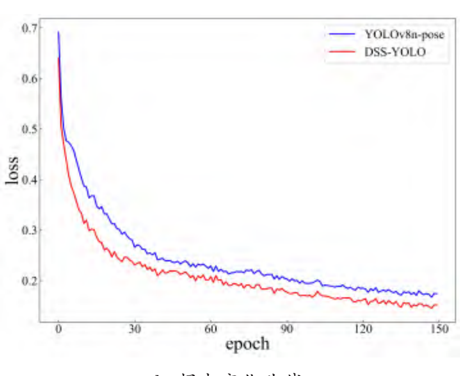
式中：i为图像序号；n为试验数据总数；$\hat{x}_i$为自动测量值；x_i为人工测量值。

2.3 模型训练及验证结果分析

为了评估改进方法优化效果，通过将训练过程可视化，对比训练过程中原模型（YOLOv8n-pose）和改进后模型（DSS-YOLO）的精度和损失变化情况，模型训练过程如图12所示。可以看出在经过150个epoch的训练周期之后两个模型精度曲线和损失曲线基本达到收敛，并且改进后的DSS-YOLO相较于原模型在训练过程中的收敛速度更快、精度更高，验证了本研究改进方法的有效性。改进对比结果总结于表2。



a. 精度变化曲线



b. 损失变化曲线

图12 YOLOv8n-pose和DSS-YOLO模型训练过程对比

Fig. 12 Comparison of YOLOv8n-pose and DSS-YOLO model training process

表2 YOLOv8n-pose和DSS-YOLO性能对比结果

Table 2 YOLOv8n-pose and DSS-YOLO performance comparison results

模型	mAP/%	参数量/×10^6 M	运算量/G
YOLOv8n-pose	88.2	3.15	8.7
DSS-YOLO	92.5	3.48	9.1

与YOLOv8n-pose相比，DSS-YOLO在关键点检测精度方面提高4.3%，同时其参数量和运算量增加相对较小。为了进一步验证本研究方法的鲁棒性和适用性，模拟在真实养殖环境中可能会碰到相机抖动和过曝情况，并观察在不同条件下测试集中的关键点检测效果，如图13所示，检测结果于表3所示。在不同条件下，DSS-YOLO均能凭借较小的计算资源完成对处于不同姿态下的蒙古马测量关键点检测，误差值均小于8.5个像素，表明模型的泛用性较强。

表3 DSS-YOLO模型对测试集中不同条件下的RGB图检测结果统计

Table 3 DSS-YOLO model detection results of RGB images under different conditions in test set

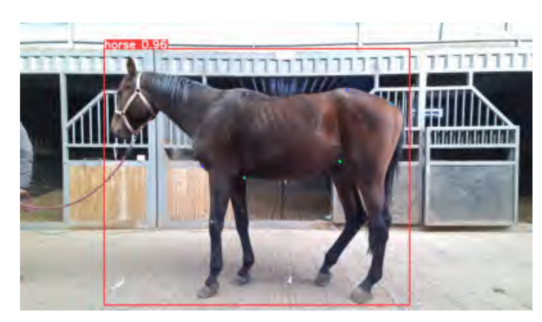
模型	条件	d_{DSS}/pixel
YOLOv8n-pose	高质量	12.4
	模糊	13.9
	曝光	15.2
DSS-YOLO	高质量	7.2
	模糊	7.9
	曝光	8.3

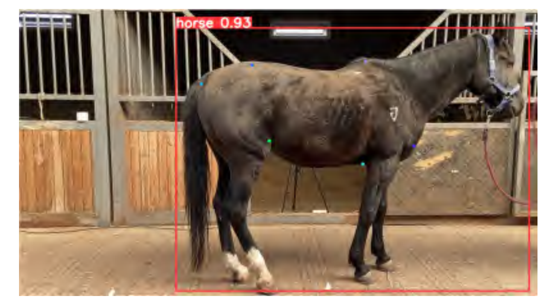
2.4 不同关键点检测模型对比试验结果

为了进一步评估本研究模型的关键点检测性能，分别采用目前主流的关键点检测模型Hourglass、HRNet和SimCC，并基于自制关键点数据集进行训练。完成训练后，模型验证最终结果如表4所示，其中DSS-YOLO的mAP值为92.5%，相比于Hourglass、HRNet和SimCC分别提高3.6%、2.8%和1.6%；d_{DSS}为7.2，相比于其余模型分别减小8.2、5.4和3.2个像素。表明DSS-YOLO模型的检测精度在验证集或测试集中都有着较高精度。在模型轻量化方面，参数量相比于其余模型分别减少91.36、25.04和23.26 M；运算量相比于Hourglass和HRNet分别减少19.6和7.7 G。

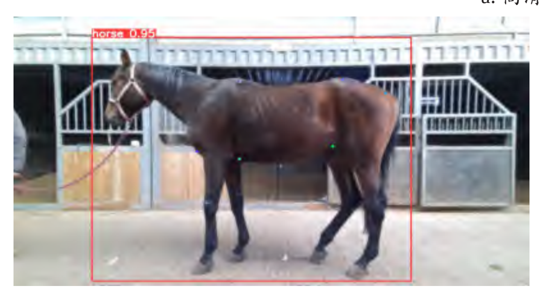
2.5 蒙古马点云重建结果分析

蒙古马点云最终重建结果如图14所示。经过一系列点云处理步骤之后，与全局点云相比，无关的背景点云和大部分噪声点云能被有效去除，同时对蒙古马三维点云中与体尺参数测量相关的信息进行分析和提取，获得了高精度的蒙古马体表点云信息。此外，通过将点云质心作为坐标系原点，设定体尺参数测量坐标系，为各项体尺参数的自动测量奠定了良好的基础。



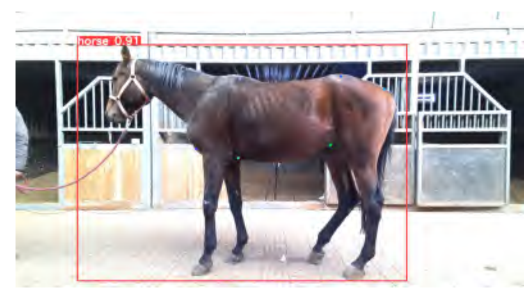


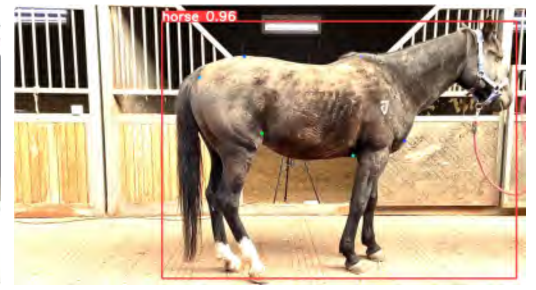
a. 高清条件





b. 模糊条件





c. 曝光条件

图 13　DSS-YOLO模型检测效果示例

Fig. 13 Example of the detection effect of DSS-YOLO model

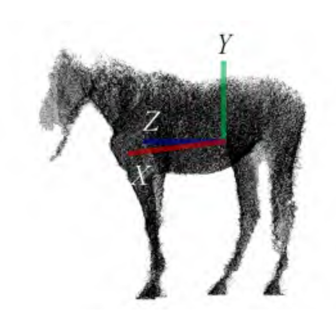
表4　不同关键点检测模型性能对比结果

Table 4 Performance comparison of different keypoint detection models

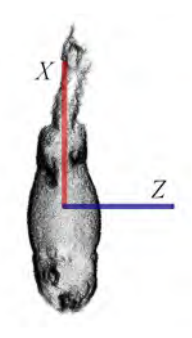
模型	mAP/%	d_{DSS}/pixel	参数量/×10^6 M	运算量/G
Hourglass	88.9	15.4	94.84	28.7
HRNet	89.7	12.6	28.52	16.8
SimCC	90.9	10.4	26.74	6.8
DSS-YOLO	92.5	7.2	3.48	9.1

2.6 体尺测量结果分析

为验证和分析本研究提出的蒙古马体尺参数自动测量方法精度，以测试集采集时的人工测量值为参考值。表5总结了各项体尺参数的实际测量值的MAE和MRE。此外，采用箱线图分析人工测量值与自动测量值的绝对误差，如图15所示。箱线图能够直观地展示测量误差的整体特征，其中箱体表示测量结果的绝对误差值分布情况；箱体的中间线代表数据中位数，即平均绝对误差值。



a. 点云斜视图



b. 点云俯视图

图 14　蒙古马点云重建结果

Fig. 14 The Mongolian horse point cloud reconstruction results

统计整体测量结果的MRE为2.29%，MAE为3.77 cm。其中，体高的MRE为1.93%，绝对误差控制为−4.41～4.43 m；体斜长的MRE为2.51%，绝对误差控制为−5.95～5.88 cm；臀高的MRE为2.02%，绝对误差控制为−4.92～4.82 cm；胸围的

表5 蒙古体尺参数测量结果平均绝对误差和平均相对误差

Table 5 The MAE and MRE of Mongolian horses body measurements

指标	体高	体斜长	臀高	胸围	臀围
MAE/cm	2.98	3.91	3.12	4.49	4.35
MRE/%	1.93	2.51	2.02	2.46	2.51

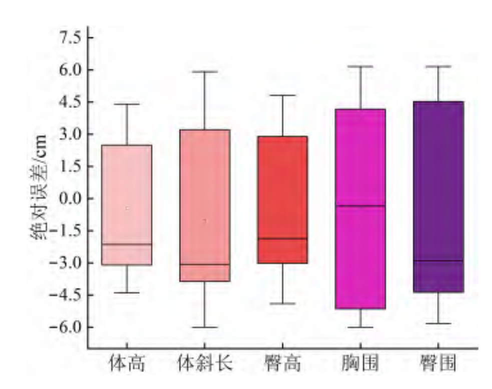


图15 各项体尺参数测量结果绝对误差箱线图

Fig. 15 Box plots of MAE for each body measurements

MRE为2.46%，绝对误差控制为−6.03～6.17 cm；臀围的MRE为2.51%，绝对误差控制为−5.83～6.16 cm。

由于蒙古马的不同体尺参数共享相同的体尺测量点，基于轮廓特征点的关键点定位方法虽然具有一定的可取性，但是该方法容易受到目标马匹姿势、养殖环境及设备精度等因素的影响，导致关键点定位不准确。本研究构建基于卷积神经网络的二维关键点检测模型，结合三维点云处理方法，可高效、准确地获取蒙古马自然状态下的体尺参数。

3 结 论

为提高蒙古马体尺参数测量效率，加快选育进程，本研究提出DSS-YOLO模型结合三维点云处理的体尺自动检测方法，结论如下。

1）本研究在YOLOv8n-pose模型的基础上，通过设计C2f_DCN模块，在Backbone网络中融合可变形卷积（DCNv2）；在Neck网络中添加SA注意力模块；将原始损失函数替换为SIoU，在模型训练过程中采用余弦退火法动态调节学习率。改进后得到的DSS-YOLO模型参数和运算量分别为3.48 M和9.1 GFLOPs，而mAP达到92.5%，在测试集中表现出较高的预测精度。结果表明，DSS-YOLO模型能满足真实场景中蒙古马测量关键点的自动检测需求。

2）完成蒙古马三维点云重建与处理分析后，基于DSS-YOLO模型的检测结果，可实现蒙古马体高、体斜长、臀高、胸围和臀围5项体尺参数的自动测量。通过在真实场景中选取30匹蒙古马进行现场测试，各项体尺参数的MRE为2.29%，MAE为3.77 cm。表明本研究所提出的方法精度较高，可满足实际生产需求。

3）本研究所提出的模型在自制数据集上表现良好，但在应对不同毛色、场景下的蒙古马可能表现不佳。因此，后续考虑使用迁移学习方法训练模型，以适应各种可能的马匹毛色和更加丰富的真实养殖环境。此外，由于本研究所使用Azure Kinect相机在某些情况下存在精度误差，后续研究考虑采用多线激光雷达结合RGB相机，并移植深度学习模型以构建蒙古马移动测量平台，满足实际操作中灵活性和简便操作的应用需要。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献：

[1] 芒来, 白东义. 马业发展 种业为先[J]. 北方经济, 2019(11): 20-25.
MANG L, BAI D Y. Analysis of the current situation of the horse industry in Inner Mongolia autonomous region[J]. Northern economy, 2019(11): 20-25.

[2] 曹晓娟, 王怀栋, 王勇. 基于SWOT分析的我国马产业发展对策[J]. 黑龙江畜牧兽医, 2020(10): 23-28.
CAO X J, WANG H D, WANG Y. Countermeasures for the development of China's horse industry based on SWOT analysis[J]. Heilongjiang animal science and veterinary medicine, 2020(10): 23-28.

[3] GHEZELSOFLOU H, HAMIDI P, GHARAHVEYSI S. Study of factors affecting the body conformation traits of Iranian Turkoman horses[J]. Journal of equine science, 2018, 29(4): 91-96.

[4] ROSENGREN M K, SIGURÐARDÓTTIR H, ERIKSSON S, et al. A QTL for conformation of back and croup influences lateral gait quality in Icelandic horses[J]. BMC genomics, 2021, 22(1): 1-13.

[5] PAKSOY Y, ÜNAL N. Multivariate analysis of morphometry effect on race performance in Thoroughbred horses[J]. Revista brasileira de zootecnia, 2019, 48: ID e20180030.

[6] DOS SANTOS M R, FREIBERGER G, BOTTIN F, et al. Evaluation of methodologies for equine biometry[J]. Livestock science, 2017, 206: 24-27.

[7] WANG Z, SHADPOUR S, CHAN E, et al. ASAS-NANP SYMPOSIUM: Applications of machine learning for livestock body weight prediction from digital images[J]. Journal of animal science, 2021, 99(2): ID skab022.

[8] DOHMEN R, CATAL C, LIU Q. Computer vision-based weight estimation of livestock: A systematic literature review[J]. New zealand journal of agricultural research, 2022, 65(2-3): 227-247.
[9] NIR O, PARMET Y, WERNER D, et al. 3D Computer-vision system for automatically estimating heifer height and body mass[J]. Biosystems engineering, 2018, 173: 4-10
[10] KRISTJANSSON T, BJORNSDOTTIR S, ALBERTSDÓTTIR E, et al. Association of conformation and riding ability in Icelandic horses[J]. Livestock science, 2016, 189: 91-101.
[11] PALLOTTINO F, STERI R, MENESATTI P, et al. Comparison between manual and stereovision body traits measurements of Lipizzan horses[J]. Computers and electronics in agriculture, 2015, 118: 408-413.
[12] GMEL A I, BURREN A, NEUDITSCHKO M. Estimates of genetic parameters for shape space data in franches-montagnes horses[J]. Animals, 2022, 12(17): ID 2186.
[13] PÉREZ-RUIZ M, TARRAT-MARTÍN D, SÁNCHEZ-GUERRERO M J, et al. Advances in horse morphometric measurements using LiDAR[J]. Computers and electronics in agriculture, 2020, 174: ID 105510.
[14] FREITAG G P, DE LIMA L G F, JACOMINI J A, et al. An accurate image analysis method for estimating body measurements in horses[J]. Journal of equine veterinary science, 2021, 101: ID 103418.
[15] LI K, TENG G. Study on body size measurement method of goat and cattle under different background based on deep learning[J]. Electronics, 2022, 11(7): ID 993.
[16] WANG X, WANG W, LU J, et al. HRST: An improved hrnet for detecting joint points of pigs[J]. Sensors, 2022, 22(19): ID 7215.
[17] ZHENG H, FANG C, ZHANG T, et al. Shank length and circumference measurement algorithm of breeder chickens based on extraction of regional key points[J]. Computers and electronics in agriculture, 2022, 197: ID 106989.
[18] 赵宇亮, 曾繁国, 贾楠, 等. 基于DeepLabCut算法的猪只体尺快速测量方法研究[J]. 农业机械学报, 2023, 54(2): 249-255, 292.
ZHAO Y L, ZENG F G, JIA N, et al. Rapid measurements of pig body size based on DeepLabCut algorithm[J]. Transactions of the Chinese society for agricultural machinery, 2023, 54(2): 249-255, 292.
[19] RUSSELL B C, TORRALBA A, MURPHY K P, et al. LabelMe: A database and web-based tool for image annotation[J]. International journal of computer vision, 2008, 77 (1): 157-173.
[20] TERVEN J, CORDOVA-ESPARZA D. A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond[EB/OL]. arXiv:2304.00501, 2023.
[21] REIS D, KUPEC J, HONG J, et al. Real-time flying object detection with YOLOv8[EB/OL]. arXiv: 2305.09972, 2023.
[22] ZHU X, HU H, LIN S, et al. Deformable convnets v2: More deformable, better results[C]// Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Piscataway, New Jersey, USA: IEEE, 2019: 9308-9316.
[23] DU X, CHENG H, MA Z, et al. DSW-YOLO: A detection method for ground-planted strawberry fruits under different occlusion levels[J]. Computers and electronics in agriculture, 2023, 214: ID 108304.
[24] ZHANG Q, YANG Y. SA-net: Shuffle attention for deep convolutional neural networks[C]// ICASSP 2021－2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, New Jersey, USA: IEEE, 2021: 2235-2239.
[25] LI X, WANG W, WU L, et al. Generalized focal loss: learning qualified and distributed bounding boxes for dense object detection[J]. Advances in neural information processing systems, 2020, 33: 21002-21012.
[26] GEVORGYAN Z. SIoU loss: More powerful learning for bounding box regression[EB/OL]. arXiv: 2205. 12740, 2022.
[27] MICROSOFT BUILD. Azure-Kinect-Sensor-SDK: K4a_transformation_color_image_to_depth_camera method [EB/OL]. [2023-11-20]. https://learn.microsoft.com/zh-cn/azure/kinect-dk/use-image-transformation, 2019.
[28] DU A, GUO H, LU J, et al. Automatic livestock body measurement based on keypoint detection with multiple depth cameras[J]. Computers and electronics in agriculture, 2022, 198: ID 107059.
[29] YIN L, ZHU J, LIU C, et al. Point cloud-based pig body size measurement featured by standard and non-standard postures[J]. Computers and electronics in agriculture, 2022, 199: ID 107135.

Automatic Measurement of Mongolian Horse Body Based on Improved YOLOv8n-pose and 3D Point Cloud Analysis

LI Minghuang[1,2], SU Lide[1,2], ZHANG Yong[1,2*], ZONG Zheying[1,2], ZHANG Shun[3]

（1. *College of Mechanical and Electronic Engineering, Inner Mongolia Agriculture University, Hohhot* 010018, *China*; 2. *Inner Mongolia Engineering Research Center of Intelligent Equipment for the Entire Process of Forage and Feed Production, Hohhot* 010018, *China*; 3. *Bayannur Modern Agriculture and Animal Husbandry Development Center of Inner Mongolia*, *Bayannur* 010051, *China*）

Abstract:

[Objective] There exist a high genetic correlation among various morphological characteristics of Mongolian horses. Utilizing ad-

vanced technology to obtain body structure parameters related to athletic performance could provide data support for breeding institutions to develop scientific breeding plans and establish the groundwork for further improvement of Mongolian horse breeds. However, traditional manual measurement methods are time-consuming, labor-intensive, and may cause certain stress responses in horses. Therefore, ensuring precise and effective measurement of Mongolian horse body dimensions is crucial for formulating early breeding plans.

[Method] Video images of 50 adult Mongolian horses in the suitable breeding stage at the Inner Mongolia Agricultural University Horse Breeding Technical Center was first collected. Fifty images per horse were captured to construct the training and validation sets, resulting in a total of 2 500 high-definition RGB images of Mongolian horses, with an equal ratio of images depicting horses in motion and at rest. To ensure the model's robustness and considering issues such as angles, lighting, and image blurring during actual image capture, a series of enhancement algorithms were applied to the original dataset, expanding the Mongolian horse image dataset to 4 000 images. The YOLOv8n-pose was employed as the foundational keypoint detection model. Through the design of the C2f_DCN module, deformable convolution (DCNV2) was integrated into the C2f module of the Backbone network to enhance the model's adaptability to different horse poses in real-world scenes. Besides, an SA attention module was added to the Neck network to improve the model's focus on critical features. The original loss function was replaced with SCYLLA-IoU (SIoU) to prioritize major image regions, and a cosine annealing method was employed to dynamically adjust the learning rate during model training. The improved model was named DSS-YOLO (DCNv2-SA-SIoU-YOLO) network model. Additionally, a test set comprising 30 RGB-D images of mature Mongolian horses was selected for constructing body dimension measurement tasks. DSS-YOLO was used for keypoint detection of body dimensions. The 2D keypoint coordinates from RGB images were fused with corresponding depth values from depth images to obtain 3D keypoint coordinates, and Mongolian horse's point cloud information was transformed. Point cloud processing and analysis were performed using pass-through filtering, random sample consensus (RANSAC) shape fitting, statistical outlier filtering, and principal component analysis (PCA) coordinate system correction. Finally, body height, body oblique length, croup height, chest circumference, and croup circumference were automatically computed based on keypoint spatial coordinates.

[Results and Discussion] The proposed DSS-YOLO model exhibited parameter and computational costs of 3.48 M and 9.1 G, respectively, with an average accuracy $mAP_{0.5:0.95}$ reaching 92.5%, and a d_{DSS} of 7.2 pixels. Compared to Hourglass, HRNet, and SimCC, $mAP_{0.5:0.95}$ increased by 3.6%, 2.8%, and 1.6%, respectively. By relying on keypoint coordinates for automatic calculation of body dimensions and suggesting the use of a mobile least squares curve fitting method to complete the horse's hip point cloud, experiments involving 30 Mongolian horses showed a mean average error (MAE) of 3.77 cm and mean relative error (MRE) of 2.29% in automatic measurements.

[Conclusions] The results of this study showed that DSS-YOLO model combined with three-dimensional point cloud processing methods can achieve automatic measurement of Mongolian horse body dimensions with high accuracy. The proposed measurement method can also be extended to different breeds of horses, providing technical support for horse breeding plans and possessing practical application value.

Key words: Mongolian horse; body measurements; convolutional neural network; attention mechanism; 3D point cloud processing; YOLOv8n-pose

Foundation items: National Natural Science Foundation of China(32360856); Inner Mongolia Autonomous Region Natural Science Foundation (2022QN03019); Science and Technology Innovation Project of Higher Education Team in Inner Mongolia Autonomous Region (NMGIRT2312); Research Initiation Program for the Introduction of High-level Talents in Inner Mongolia Agricultural University (NDYB2020-21)

Biography: LI Minghuang, E-mail: a1286862045@163.com

***Corresponding author:** ZHANG Yong, E-mail: yongz@imau.edu.cn